Hi Sandy,

First off, thanks for taking the time to speak to me.

As I mentioned on the phone, we were initially looking at starting up a
"data mining" and "search indexing" service for TV broadcast (and
cable and satellite, but I'll group them all together) news services...
The idea was to provide a "Google.com"-like service where someone could
type in a phrase, and the transcripts of all indexed (recorded) news
programmes would be searched in much the same way... Allowing one
to find mention of certain current events and topics in TV news exactly
as one might "google" for a web page.

Unfortunately, testing our service showed that while it was a good idea,
there were too many problems with the data that we were getting from
the networks: "garbage in, garbage out" as they say.

Comparing tapes of news broadcasts against the transcriptions (or
simply reading along the closed-captions while watching news) you can
see exactly what we found:

* proper names (names of places and peoples) were often transcribed
  incorrectly. For example, on Fox News, the battle of "el Shakkar" was
  captioned as "Al Shaker".

* sometimes the caption transcriber would struggle with a word or name,
  and while they were trying to figure out what it had been, they would
  fall behind... and when they were finally ready, they would have omitted
  one or more phrases or sentences just to get resynchronized with the
  soundtrack.

* scripted material would sometimes be edited or even improvised, such

that while the closed captioning was what had originally been
scripted, it would be very different from what was eventually
ad-libbed on the air.

* sometimes editing changes (due to inability to find footage in time,
  etc) would cause the sequencing of reporting to vary from the
  originally scripted (and transcribed order), such that the order of
  various topics would not be captioned (text) in the same sequence that
  they were seen (video) and heard (audio).

These were just some of the more predominant issues. There were various
other issues that were either less frequent, or else we couldn't readily
divine their origins of such problems.

In any case, we drew two conclusions.

(1) while the idea for the service was sound and viable, the quality of
    the closed captions we were receiving would make the service too
    unreliable (high rate of inaccurate or missing "hits" to the source
    material);

(2) but in a broader sense, we were glad that we all had more or less
    adequate hearing and we're relying on the quality of closed captions
    in TV news to be informed (and, jokingly, not having been able to
    buy those 3rd row seats to a very loud Rolling Stones concert in
    1982 was in retrospect a blessing).

So I decided to contact your office and raise the issue of quality
metrics for closed-captions in TV content (and journalism/documentaries
in particular).

We'd like to see an industry-wide push for better quality of closed
captioning, and perhaps independent monitoring of closed captioning
with fixed benchmarks for measuring their accuracy.

That is, the Department of Education or the FCC could establish metrics
whereby the accuracy and completeness of closed captions could be
objectively measured (perhaps using a point system for misspellings,
omissions, etc) that would be a criteria in license renewal for a

broadcaster (or content provider). A certain "passing grade" would be required to maintain one's license.

Or, alternatively, a cooperative strategy could be used with industry to have them establish their own watchdog group and develop their own grading system for acceptable levels of service.

Sometimes the latter approach works. Sometimes it doesn't.

These are a few of my suggestions and my understanding of the problem and its likely causes.

I'd be interested to know what the Senator's office view is on this issue.

Sincerely,

Philip A. Prindeville
Boise, ID